

The Shapley Value in Knapsack Budgeted Games

Smriti Bhagat¹, Anthony Kim^{2*}, S. Muthukrishnan³, and Udi Weinsberg¹

¹ Technicolor Research

{smriti.bhagat,udi.weinsberg}@technicolor.com

² Department of Computer Science, Stanford University

tonyekim@stanford.edu

³ Rutgers University

muthu@cs.rutgers.edu

Abstract. We propose the study of computing the Shapley value for a new class of cooperative games that we call budgeted games, and investigate in particular knapsack budgeted games, a version modeled after the classical knapsack problem. In these games, the “value” of a set S of agents is determined only by a critical subset $T \subseteq S$ of the agents and not the entirety of S due to a budget constraint that limits how large T can be. We show that the Shapley value can be computed in time faster than by the naïve exponential time algorithm when there are sufficiently many agents, and also provide an algorithm that approximates the Shapley value within an additive error. For a related budgeted game associated with a greedy heuristic, we show that the Shapley value can be computed in pseudo-polynomial time. Furthermore, we generalize our proof techniques and propose what we term algorithmic representation framework that captures a broad class of cooperative games with the property of efficient computation of the Shapley value. The main idea is that the problem of determining the efficient computation can be reduced to that of finding an alternative representation of the games and an associated algorithm for computing the underlying value function with small time and space complexities in the representation size.

1 Introduction

The Shapley value is a well-studied solution concept for fair distribution of profit among agents in cooperative game theory. Given a coalition of agents that collectively generate some profit, fair distribution is important to maintain a stable coalition such that no subgroup of agents has an incentive to unilaterally deviate and form its own coalition. While the Shapley value is not a stability concept, it uniquely satisfies a set of desirable properties for fair profit distribution based on individual contributions. It has been shown useful on a wide range of cooperative games and, more recently, applied beyond the game-theoretic setting in problems related to social networks [18, 16] and computer networks [13, 17].

* This work was done while the author was an intern at Technicolor Research Lab. Supported in part by an NSF Graduate Research Fellowship.

Efficient — (pseudo) polynomial time — computation of the Shapley value has been studied for many classes of cooperative games. One such example is weighted voting games that model parliamentary voting where agents are parties, the weight of each party is the number of the same party representatives, and a coalition of parties is winning (has value 1) if its total weight is at least some quota, or losing (has value 0) otherwise. It was shown that computing the Shapley value in the weighted majority games, where the quota is half the total weight of all the agents, is #P-complete [7] and NP-hard [15]. Note, however, that there is a pseudo-polynomial time algorithm using dynamic programming [14].

In another line of research, representation schemes for cooperative games have been proposed in [6, 10, 11, 1]; if a given cooperative game has a small alternative representation in one of these schemes, then the Shapley value can be computed efficiently in time polynomial in the size of the alternative representation. For example, we can represent a given cooperative game as a collection of smaller cooperative games in multi-issue representation [6], or in terms of logic rules in marginal contribution net representation [10].

We propose a new class of cooperative games that we call *budgeted games* and study the Shapley value computation in these games. In cooperative games, the value function $v(S)$ for a coalition S is determined by all the agents in S , but may explicitly depend on a sub-coalition in some domains (e.g., [3, 5, 2]). We study value functions conditioned on a budget B where $v(S)$ may be totally determined by a potentially strict subset $T \subset S$ of agents. That is, budget B models a physical or budget constraint that may limit the actual value of a coalition to be less than simply the total aggregate value of all the individual contributions and, hence, the profit generation of a coalition is determined only by a sub-coalition of the agents. There are many examples we can readily formulate as budgeted games to model real-life scenarios:

- (*Graph Problems*) Consider a network of nodes that correspond to facilities and edges between them that correspond to communication links. This can be modeled as a graph G with weights on nodes. For any subset S of nodes, $v_B(S)$, the value created by set S under budget B , may be the maximum weight of an independent set of at most B nodes.
- (*Set Problems*) Let each agent be a sales agent targeting a specific set of customers. Then $v_B(S)$ may be the maximum number of customers that can be targeted by a subset of size at most B of sales agents from S .
- (*Packing Problems*) Consider creating a task force from a pool S of available agents where each agent is associated with some value and cost. Then $v_B(S)$ may be the largest total aggregate value from a subset of the agents with total cost at most B .
- (*Data Mining Problems*⁴) Let each agent represent a document with some quality measure with respect to a fixed search query. We may approximate the total value of an ordered list of documents S , ordered by the quality

⁴ This is the Shapley value computation problem for what is commonly known as the Top- k problem.

measure, by those that appear at the top of the list. Then, the corresponding $v_B(S)$ is the sum of the top B quality scores of documents in S .

For the Shapley value to be useful in value division problems modeled as budgeted games, we cannot simply apply the formula for the Shapley value as it would lead to an exponential time algorithm. Hence, it is important to understand its computational complexity in these games, and we study the knapsack version (equivalently, Packing Problems) in this paper. As far as we know, the budgeted games have not been studied previously. A related class of games called bin-packing games [8, 12, 19] has been studied for different solution concepts of core and ϵ -core.⁵

Our Contributions. First, we propose a new class of cooperative games, *budgeted games*, and investigate the computational complexity of the Shapley value in a particular version of budgeted games. Second, we generalize our proof techniques and propose a general framework, *algorithmic representation*, for cooperative games. We note that all our algorithms have running times with a polynomial dependence on the number of agents. More specifically, our contributions are as follows:

- We study the knapsack version of budgeted games and show that computing the Shapley value in these games is NP-hard. On the other hand, we show that the Shapley value can be computed in time faster than by the naïve exponential time algorithm when there are sufficiently many agents.
- We provide an additive approximation scheme for the Shapley value via rounding; our approach does not use the standard sampling and normal distribution techniques [4, 9] in estimating the Shapley value.
- We consider the value function obtained by a 2-approximation greedy algorithm for the classical knapsack problem and show that for this function, the Shapley value can be computed in pseudo-polynomial time.
- We provide generalizations and present the algorithmic representation framework that captures a broad class of cooperative games with the property of efficient computation of the Shapley value. This includes many known classes of cooperative games in [7, 14, 16] and those with concise representations using schemes in [6, 10, 1].

2 Preliminaries

We represent the profit distribution problem as a *cooperative game* (N, v) where N is the set of agents and $v : 2^N \rightarrow \mathbb{R}$ is the *characteristic function* that assigns a value to each subset of agents, with $v(\emptyset) = 0$. We also call v the *value function* and use both characteristic and value functions interchangeably. For a subset of

⁵ While items and bins separate and bins model linear constraints in knapsack budgeted games, both items and bins are treated as agents and the goal is to share profit among them in a fair way in bin-packing games.

agents $S \subseteq N$, we interpret $v(S)$ as the value that these agents can generate collectively; $v(N)$ is the total value that the whole group generates.

The *Shapley value* [20] is a solution concept based on marginal contributions that divides the total value $v(N)$ into individual shares $\phi_1, \dots, \phi_{|N|}$ satisfying an intuitive notion of fairness. For $i \in N$ and $S \subseteq N \setminus \{i\}$, we define agent i 's *marginal contribution to S* to be $v(S \cup \{i\}) - v(S)$. The Shapley value is the unique profit distribution solution that satisfies the following properties:

1. (Efficiency) $\sum_{i \in N} \phi_i(v) = v(N)$;
2. (Symmetry) If $v(S \cup \{i\}) - v(S) = v(S \cup \{j\}) - v(S)$ for all $S \subseteq N \setminus \{i, j\}$, then $\phi_i(v) = \phi_j(v)$;
3. (Null Player) If $v(S \cup \{i\}) - v(S) = 0$ for all $S \subseteq N \setminus \{i\}$, then $\phi_i(v) = 0$;
4. (Linearity) For any two cooperative games (N, v) and (N, w) and their combined game $(N, v + w)$, $\phi_i(v) + \phi_i(w) = \phi_i(v + w)$ for all $i \in N$.

The Shapley value for each agent i is computed as

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S)). \quad (1)$$

Note the Shapley value is a weighted average of agent i 's marginal contributions. Equivalently, it can also be computed as $\phi_i(v) = \frac{1}{|N|!} \sum_{\pi \in \Pi} v(P_\pi^i \cup \{i\}) - v(P_\pi^i)$, where Π is the set of all $|N|!$ permutations of the agents and P_π^i is the set of agents preceding agent i in the order represented by permutation π .

There are two sources of computational complexity in the Shapley value: an exponential number of terms in the summation and individual evaluations of the characteristic function v . Directly applying the above equations leads to a naïve algorithm with running time at least exponential in the number of agents, $\Omega(2^{|N|})$; furthermore, each individual evaluation of v can be expensive.

3 Knapsack Budgeted Games

A *knapsack budgeted game* (N, v) is a cooperative game with the alternative representation given by a nonnegative integer tuple $(\{(l_1, w_1), \dots, (l_{|N|}, w_{|N|})\}, l_{\text{bin}})$ such that $v(S) = \max_{S' \subseteq S: l(S') \leq l_{\text{bin}}} w(S')$ for all $S \subseteq N$, where $l(S') = \sum_{k \in S'} l_k$ and $w(S') = \sum_{k \in S'} w_k$. Each agent i is described by (l_i, w_i) where l_i and w_i are the agent's length and weight, respectively. The variable l_{bin} is the bin size that restricts which set of agents can directly determine the value function v . For a set of agents S , the value $v(S)$ is determined by solving an optimization problem where the total value of selected agents, possibly a strict subset of S , is optimized subject to a budget constraint; the other unselected agents do not contribute explicitly. Note the similarities with the classical knapsack problem in which the objective is to find the maximum total value of items that can be packed into a fixed size bin.

Knapsack budgeted games are useful when the characteristic function v of a cooperative game can be modeled as the objective value of an optimization problem subject to linear constraints. In this paper, we only consider the games with a

single linear constraint, but our results extend to knapsack budgeted games with multiple linear constraints. In a knapsack budgeted game with multiple budget constraints, each agent is associated with a length vector $\mathbf{l} = (l^1, \dots, l^d)$ and a weight and there is a budget constraint on each coordinate, i.e., $l_{\text{bin}}^1, \dots, l_{\text{bin}}^d$, assuming d budget constraints.

For an application, we can use knapsack budgeted games and the Shapley value to model value division in a sport team. We would like to give out bonuses proportional to the Shapley value solution. Assume each player i is associated with a skill level w_i and, in a game of the sport, at most B players from each team can play. We model the value of the team as the total aggregate skill level of its best B players, since they usually start and play the majority of the games. Then, this is a knapsack budgeted game with skill levels as weights, unit lengths, and $l_{\text{bin}} = B$. Note the Shapley value of a player not in the top B may be positive. Since he is still contributing to the team as a reserve player and might be one of the top B players in a subset of the team, say available players in an event of injury, he should be compensated accordingly.

In the following sections, we assume that the knapsack budgeted game (N, v) has the representation $(\{(l_1, w_1), \dots, (l_{|N|}, w_{|N|})\}, l_{\text{bin}})$. We define $w_{\max} = \lceil l_{\text{bin}} \cdot \max_i w_i / l_i \rceil$, which is an upper bound on the value $v(N)$. We use shorthand notations $l(S) = \sum_{k \in S} l_k$ and $w(S) = \sum_{k \in S} w_k$ for any subset S . The set of agents are ordered and labeled with $1, \dots, |N|$. For a set of agents X and two integers a and b , we use $X_{a,b}$ to denote the subset $\{i \in X : a \leq i \leq b\}$. To avoid degenerate cases, we further assume $0 < l_i \leq l_{\text{bin}}$ for all i . We use the indicator function Id that equals to 1 if all the input conditions hold, or 0 otherwise.

4 The Shapley Value in Knapsack Budgeted Games

We present a hardness result, an algorithm for computing the Shapley value exactly, and a deterministic approximation scheme that approximates within an additive error.

4.1 Exact Computation

By the NP-completeness of the classical knapsack problem and the efficiency property of the Shapley value, it follows that (see Appendix A for details):

Theorem 1. *The problem of computing the Shapley value in the knapsack budgeted games is NP-hard.*

While a polynomial time algorithm for computing the Shapley value may or may not exist, the naïve exponential time algorithm is too slow when $|N|$ is large. When $|N|$ is sufficiently large, especially when $|N| \gg l_{\text{bin}}$, we show that a faster algorithm exists:

Theorem 2. *In the knapsack budgeted games, the Shapley value can be computed in time $O(l_{\text{bin}}(w_{\max} + 1)^{l_{\text{bin}}+1}|N|^2)$ for each agent.*

To prove Theorem 2, we associate each subset $S \subseteq N$ with a vector from a finite-sized vector space that completely determines an agent's marginal contribution to S . If the cardinality of the vector space is small and the partitions of the $2^{|N|}$ subsets corresponding to the vectors can be found efficiently, we can evaluate v once for each vector instead of once for each subset, reducing the overall computation time. Note that the well-known dynamic programming algorithm, call it \mathcal{A} , for the classical knapsack problem can be used to compute v ; for a given S , the algorithm iteratively updates an integer array of length $l_{\text{bin}} + 1$ holding the optimal values for the sub-problems with smaller bin sizes and returns a final value determined by the array at termination.⁶ We associate with each subset S the final state of the array when \mathcal{A} runs on S and determine the cardinalities of resulting partitions using a dynamic program, different but related to \mathcal{A} ; the dynamic program counts the number of optimal solutions to the sub-problems grouped by objective value while \mathcal{A} simply computes the optimal solutions to the sub-problems.

We use the following lemma to prove Theorem 2; it shows that if the set of possible marginal contribution values for agent i is small, then we can reduce the number of evaluations of v by grouping subsets of $N \setminus \{i\}$ by marginal contribution value and evaluating v once for each group (see Appendix A for a proof).

Lemma 1. *Assume there exist positive integers p_i and partition functions $P_i : 2^{N \setminus \{i\}} \rightarrow \{1, \dots, p_i\}$, for $i = 1, \dots, |N|$, such that if $P_i(S) = P_i(S')$ for two different $S, S' \subseteq N \setminus \{i\}$, then $v(S \cup \{i\}) - v(S) = v(S' \cup \{i\}) - v(S')$. Let $m_i(p)$ be agent i 's marginal contribution to S for all S satisfying $P_i(S) = p$, and $c(i, s, p) = \#\{S \subseteq N \setminus \{i\} : |S| = s, P_i(S) = p\}$ for $i \in N$, $0 \leq s \leq |N| - 1$, and $1 \leq p \leq p_i$. Then, the Shapley value for agent i can be computed as*

$$\phi_i = \sum_{p=1}^{p_i} \sum_{s=0}^{|N|-1} c(i, s, p) \frac{s!(|N|-s-1)!}{|N|!} m_i(p)$$

in time $O(p_{\max}(t + q)|N|)$, where $p_{\max} = \max_i p_i$, t is an upper bound on the computation time of the coefficients c , and q is the evaluation time of v .

We now prove Theorem 2 by applying Lemma 1:

Proof. (of Theorem 2) We compute the Shapley value for some fixed agent i . We define $V_{A,b} = \max_{S' \subseteq A : l(S') \leq b} w(S')$, for $A \subseteq N$ and $0 \leq b \leq l_{\text{bin}}$, and vector $\mathbf{V}_S = (V_{S,0}, \dots, V_{S,l_{\text{bin}}})$, for subsets $S \subseteq N$. Let \mathcal{V} be the finite vector space $\{0, \dots, w_{\max}\}^{l_{\text{bin}}+1}$ that contains vectors \mathbf{V}_S . We use the 0-based index to indicate coordinates of a vector in \mathcal{V} ; so, $v(S) = V_{S,l_{\text{bin}}} = \mathbf{V}_S(l_{\text{bin}})$ for all S . Given \mathbf{V}_S , agent i 's marginal contribution to S can be computed in constant time as $v(S \cup \{i\}) - v(S) = \max\{\mathbf{V}_S(l_{\text{bin}} - l_i) + w_i - \mathbf{V}_S(l_{\text{bin}}), 0\}$. Let this expression

⁶ Assume the agents in S are labeled $1, \dots, |S|$ for simplicity. For $1 \leq j \leq |S|$, we define $c(j, b) = \max_{S' \subseteq S_{1,j} : l(S') \leq b} w(S')$. It has the recurrence relation $c(j, b) = \max\{c(j-1, b), c(j-1, b-l_j) + w_j\}$. We compute $c(j, b)$'s and, hence, $v(S) = c(|S|, l_{\text{bin}})$ in $O(|S|l_{\text{bin}})$ time.

be defined more generally as $m_i(\mathbf{v}) = \max\{\mathbf{v}(l_{\text{bin}} - l_i) + w_i - \mathbf{v}(l_{\text{bin}}), 0\}$ for $\mathbf{v} \in \mathcal{V}$.

We partition $2^{N \setminus \{i\}}$ by the pair $(|S|, \mathbf{V}_S)$ so that for each possible (s, \mathbf{v}) pair, all subsets S satisfying $|S| = s$ and $\mathbf{V}_S = \mathbf{v}$ are grouped together. Clearly, the marginal contribution of agent i is the same within each partition. To compute the cardinality of each partition, we use dynamic programming. Let $N' = N \setminus \{i\}$, ordered and relabeled $1, \dots, |N| - 1$. For $0 \leq j \leq |N| - 1$, $0 \leq s \leq j$, and $\mathbf{v} \in \mathcal{V}$, we define $\hat{c}(j, s, \mathbf{v}) = \#\{S \subseteq N'_{1,j} : |S| = s, \mathbf{V}_S = \mathbf{v}\}$. Note \hat{c} has the recurrence relation

$$\hat{c}(j, s, \mathbf{v}) = \hat{c}(j-1, s, \mathbf{v}) + \sum_{\mathbf{u}: \text{UPDATE}(\mathbf{u}, l'_j, w'_j) = \mathbf{v}} \hat{c}(j-1, s-1, \mathbf{u}),$$

with the base case $\hat{c}(0, 0, \mathbf{0}) = 1$, where l'_j and w'_j correspond to the j -th agent in order in N' and UPDATE is an $O(l_{\text{bin}})$ algorithm that updates \mathbf{u} with the additional agent: 1) Initialize $\mathbf{v} = \mathbf{u}$; 2) For $j = l'_j, \dots, l_{\text{bin}}$, $\mathbf{v}(j) = \max\{\mathbf{v}(j), \mathbf{u}(j - l'_j) + w'_j\}$; and 3) Return \mathbf{v} .

Using the recurrence relation, we compute $\hat{c}(j, s, \mathbf{v})$ for all j, s , and \mathbf{v} in time $O(l_{\text{bin}}(w_{\text{max}} + 1)^{l_{\text{bin}}+1}|N|^2)$. By Lemma 1,

$$\phi_i = \sum_{\mathbf{v} \in \mathcal{V}} \sum_{s=0}^{|N|-1} \hat{c}(|N| - 1, s, \mathbf{v}) \frac{s!(|N| - s - 1)!}{|N|!} m_i(\mathbf{v}),$$

and the Shapley value can be calculated in time $O((w_{\text{max}} + 1)^{l_{\text{bin}}+1}|N|)$ using the precomputed values of \hat{c} . The computation of \hat{c} dominates the application of the Shapley value equation, and the overall running time is $O(l_{\text{bin}}(w_{\text{max}} + 1)^{l_{\text{bin}}+1}|N|^2)$ per agent. \square

4.2 Additive Approximation

Similar to the fully polynomial time approximation scheme for the classical knapsack problem (see [21]), we show an approximation scheme for the Shapley value by rounding down the weights w_i 's and computing the Shapley value of the cooperative game (N, v') where v' is an approximation of v . Our technique of computing the Shapley value by approximating the characteristic function v is deterministic and does not require concentration inequalities like the standard statistical methods of sampling and normal distribution techniques in [4, 9].

The following lemma formalizes how an approximation of the characteristic function v leads to an additive error in the Shapley value computation (see Appendix A for a proof):

Lemma 2. *If v' is an α -additive approximation of v , i.e., $v'(S) \leq v(S) \leq v'(S) + \alpha$ for all $S \subseteq N$, then the Shapley value ϕ'_i computed with respect to v' is within an α -additive error of the Shapley value ϕ_i computed with respect to v , for all i .*

When w_{max} is sufficiently larger than l_{bin} , the approximation scheme's running time is faster than that of the exact algorithm of Theorem 2:

Theorem 3. *In the knapsack budgeted games, the Shapley value can be computed within an ϵw_{\max} -additive error in $O((l_{\text{bin}}^2/\epsilon + 1)^{l_{\text{bin}}+1}|N|^2)$ for each agent, where $\epsilon > 0$.⁷*

Proof. We construct an approximate characteristic function v' of v as follows. Let $\epsilon > 0$ and $k = \epsilon w_{\max}/l_{\text{bin}}$. Note that when $l_{\text{bin}}^2/\epsilon < w_{\max}$, $k > 1$. For each agent i , let the rounded weight w'_i be $\lfloor \frac{w_i}{k} \rfloor$. The lengths do not change. To compute $v'(S)$, we compute the optimal set $S' \subseteq S$, using dynamic programming, with respect to the rounded weights $w'_1, \dots, w'_{|N|}$ and let $v'(S) = k \sum_{i \in S'} w'_i$. In other words, $v'(S) = k \cdot \max_{S' \subseteq S: l(S') \leq l_{\text{bin}}} w'(S')$ for all $S \subseteq N$, where we use the shorthand notation $w'(S) = \sum_{k \in S} w'_k$.

We show $v(S) \geq v'(S) \geq v(S) - \epsilon w_{\max}$, for all $S \subseteq N$. Let S be a subset and $T_O, T' \subseteq S$ be the optimal subsets using original and rounded weights, respectively, such that $v(S) = w(T_O)$ and $v'(S) = k \cdot w'(T')$. Note that both optimal sets have cardinality at most l_{bin} . Because of rounding down, $w_i - kw'_i \leq k$ and $\sum_{j \in T_O} w_j - k \sum_{j \in T_O} w'_j \leq kl_{\text{bin}}$. Since T' is optimal with respect to the rounded weights, $\sum_{j \in T'} w'_j \geq \sum_{j \in T_O} w'_j$. Then, $v'(S) = k \sum_{j \in T'} w'_j \geq k \sum_{j \in T_O} w'_j \geq \sum_{j \in T_O} w_j - kl_{\text{bin}} = v(S) - \epsilon w_{\max}$. Since $w_i \geq kw'_i$ for all i , $v(S) = w(T_O) \geq w(T') \geq kw'(T') = v'(S)$. Hence, v' is an ϵw_{\max} -additive approximation of v . Then, the Shapley value computed with respect to v' is within ϵw_{\max} of the original Shapley value by Lemma 2.

We now compute the Shapley value with respect to v' . For $A \subseteq N$, $0 \leq b \leq l_{\text{bin}}$, we define $V'_{A,b} = \max_{S' \subseteq A: l(S') \leq b} w'(S')$. For a subset $S \subseteq N$, we define vector $\mathbf{V}'_S = (V'_{S,0}, \dots, V'_{S,l_{\text{bin}}})$. Note that $w'_i = \lfloor \frac{w_i}{k} \rfloor \leq \lfloor \frac{w_{\max}}{k} \rfloor = \lfloor \frac{l_{\text{bin}}}{\epsilon} \rfloor$. Then, we can upper bound $w'(S) \leq l_{\text{bin}} \lfloor \frac{l_{\text{bin}}}{\epsilon} \rfloor$, for all S . Let $\mathcal{V}' = \{0, \dots, l_{\text{bin}} \lfloor \frac{l_{\text{bin}}}{\epsilon} \rfloor\}^{l_{\text{bin}}+1}$ that vectors \mathbf{V}'_S are contained in. Note $v'(S) = k \cdot \mathbf{V}'_S(l_{\text{bin}})$ for all S . From vector \mathbf{V}'_S , we can compute agent i 's marginal contribution to S with respect to v' in constant time: $v'(S \cup \{i\}) - v'(S) = k \cdot \max\{\mathbf{V}'_S(l_{\text{bin}} - l_i) + w'_i - \mathbf{V}'_S(l_{\text{bin}}), 0\}$.

From here, we follow the proof of Theorem 2. We compute the analogue of \hat{c} in $O((l_{\text{bin}}^2/\epsilon + 1)^{l_{\text{bin}}+1}|N|^2)$, and this is the dominating term in the Shapley value computation with respect to v' . \square

5 Greedy Knapsack Budgeted Games

Motivated by the approximation scheme in Theorem 3, we investigate *greedy knapsack budgeted games*, a variant of knapsack budgeted games, and show the Shapley value in these games can be computed in pseudo-polynomial time. A greedy knapsack budgeted game has the same representation as the knapsack budgeted games, but its characteristic function is computed by a 2-approximation heuristic for the classical knapsack problem. We defer proofs to Appendix B.

⁷ For agent i , its Shapley value ϕ_i is clearly in $[0, w_{\max}]$. Using the approximation scheme, we can compute ϕ_i within $\frac{1}{7}w_{\max}$ for instance. As long as $\epsilon > l_{\text{bin}}^2/w_{\max}$, the approximation scheme has a faster running time than the exact algorithm in Theorem 2; this observation about ϵ is also true for the fully polynomial time approximation scheme for the classical knapsack problem (see [21]).

Algorithm 1 Greedy Heuristic $\mathcal{A}'(S, l_{\text{bin}})$

-
- 1: Let $a = \operatorname{argmax}_{k \in S} w_k$.
 - 2: Select agents in S in decreasing order of $\frac{w_i}{l_i}$ and stop when the next agent does not fit into the bin of size l_{bin} ; let S' be the selected agents.
 - 3: Return S' if $w(S') \geq w_a$, or $\{a\}$ otherwise.
-

Theorem 4. *In the greedy knapsack budgeted games (N, v) with $v(S) = \mathcal{A}'(S, l_{\text{bin}})$ for all S , the Shapley value can be computed in $O(l_{\text{bin}}^5 w_{\max}^5 |N|^8)$ for each agent, where the greedy heuristic $\mathcal{A}'(S, l_{\text{bin}})$ is computed as in Algorithm 1.*

While motivated by knapsack budgeted games, we use a different proof technique using the following lemma to prove Theorem 4. It generalizes the observation that in the simple cooperative game (N, v) where the agents have weights $w_1, \dots, w_{|N|}$ and the characteristic function v is additive, i.e., $v(S) = \sum_{k \in S} w_k$, the Shapley value ϕ_i is exactly w_i for all i .

Lemma 3. *Assume that the cooperative game (N, v) has a representation (M, w, A) where M is a set, $w : M \rightarrow \mathbb{R}$ is a weight function, and $A : 2^N \rightarrow 2^M$ is a mapping such that $v(S) = \sum_{e \in A(S)} w(e)$, $\forall S \subseteq N$. Let $c_+(i, s, e) = \#\{S \subseteq N \setminus \{i\} : |S| = s, e \in A(S \cup \{i\})\}$ and $c_-(i, s, e) = \#\{S \subseteq N \setminus \{i\} : |S| = s, e \in A(S)\}$, for $i \in N$, $e \in M$, and $0 \leq s \leq |N| - 1$. Then, the Shapley value for agent i can be computed as*

$$\phi_i = \sum_{e \in M} \sum_{s=0}^{|N|-1} (c_+(i, s, e) - c_-(i, s, e)) \frac{s!(|N|-s-1)!}{|N|!} w(e).$$

in time $O(t|M||N|)$ where t is an upper bound on the computation time of the coefficients c_+ and c_- .

6 Generalizations

We present generalizations of our proof techniques and propose an unifying framework that captures a broad class of cooperative games in which computing the Shapley value is tractable, including many known classes of cooperative games in [7, 14, 16] and those with concise representations using schemes in [6, 10, 1]. The main idea is that the problem of computing the Shapley value reduces to that of finding an efficient algorithm for the cooperative game's characteristic function. More precisely, if a cooperative game (N, v) is described in terms of an alternative representation I and an algorithm A with low time and space complexities that computes v , formalized in terms of *decomposition*, then we can compute the Shapley value efficiently. To illustrate the generalizations' applicability, we use them to give examples of cooperative games in which the Shapley value can be computed efficiently.

For each generalization, we consider two cases: the order-agnostic case in which A processes agents in an arbitrary order, and the order-specific case in which A processes in a specific order, like the greedy heuristic in Theorem 4.

Algorithm 2 Computing $A(I, S)$ with a decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$

```

1:  $A_{\text{setup}}(I)$  outputs  $I'$ ,  $\mathbf{x}$ 
2: for  $i \in S$  do
3:    $\mathbf{x} = A_{\text{update}}(I', i, \mathbf{x})$ 
4: end for
5: Return  $A_{\text{final}}(I', \mathbf{x})$ 

```

Definition 1. Assume a cooperative game (N, v) has an alternative representation I and a deterministic algorithm A such that $v(S) = A(I, S)$ for all $S \subseteq N$. Algorithm A has a decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$ if $A(I, S)$ can be computed as in Algorithm 2. We denote the running times of the sub-algorithms of the decomposition t_{setup} , t_{update} and t_{final} , respectively.

In Algorithm 2, $\mathbf{x} = (x_1, x_2, \dots)$ is a vector of variables that is initialized to some values independent of subset S and determines the algorithm A 's final return value. I' is an auxiliary data structure or states that only depend on the representation I and is used in subsequent steps for ease of computation; I' can be simply I if no such preprocessing is necessary. Theorem 2 can be generalized as follows:

Theorem 5. Assume a cooperative game (N, v) has an alternative representation I and a deterministic algorithm A that computes v . If A has a decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$ such that at most $n(I)$ variables \mathbf{x} are used with each taking at most $m(I)$ possible values as S ranges over all subsets of N , then the Shapley value can be computed in $O(t_{\text{setup}} + t_{\text{update}}m^n|N|^2 + t_{\text{final}}m^n|N|)$ for each agent. In order-specific cases, for Steps 2-4 of Algorithm 2, the running time is $O(t_{\text{setup}} + t_{\text{update}}m^{2n}|N|^2 + t_{\text{final}}m^{2n}|N|)$. Note that n and m are representation-dependent numbers and the argument I has been omitted.

Proof. Given the alternative representation I , we compute the Shapley value of agent i . We associate $v(S)$ with the final values, $\mathbf{x}_{S, \text{final}}$, of $n(I)$ variables \mathbf{x} in $A(I, S)$, for all $S \subseteq N \setminus \{i\}$. We partition $2^{N \setminus \{i\}}$ by the pair $(|S|, \mathbf{x}_{S, \text{final}})$ into at most $m^n|N|$ partitions, omitting the argument I from n and m . Let \mathcal{X} be the set of all possible final values of the variables \mathbf{x} ; note that its cardinality is at most m^n . We compute the cardinalities of the partitions using dynamic programming. Let $N' = N \setminus \{i\}$, ordered and relabeled $1, \dots, |N| - 1$, and $i = |N|$. For $0 \leq j \leq |N| - 1$, $0 \leq s \leq j$, and $\mathbf{v} \in \mathcal{X}$, we define $\hat{c}(j, s, \mathbf{v}) = \# \{S \subseteq N'_{1,j} : |S| = s, \mathbf{x}_{S, \text{final}} = \mathbf{v}\}$. Then, \hat{c} has the recurrence relation

$$\hat{c}(j, s, \mathbf{v}) = \hat{c}(j-1, s, \mathbf{v}) + \sum_{\mathbf{u}: A_{\text{update}}(I', j, \mathbf{u}) = \mathbf{v}} \hat{c}(j-1, s-1, \mathbf{u})$$

with the base case $\hat{c}(0, 0, \mathbf{s}) = 1$, where \mathbf{s} is the initial states of variables \mathbf{x} . Using A_{setup} , we compute I' and the initial values \mathbf{s} in $O(t_{\text{setup}})$. Using the recurrence relation and A_{update} , we compute $\hat{c}(j, s, \mathbf{v})$ for all j, s , and \mathbf{v} in time $O(t_{\text{update}}m^n|N|^2)$. Note that for a subset $S \subseteq N \setminus \{i\}$, we can compute agent i 's marginal contribution to S , i.e., $v(S \cup \{i\}) - v(S)$, in $O(t_{\text{update}} + t_{\text{final}})$ from the

Algorithm 3 Computing $A(S)$ with a per-element decomposition $\{(A_{\text{setup}}^e, A_{\text{update}}^e, A_{\text{final}}^e)\}_{e \in M}$

```

1: Initialize  $S' = \emptyset$ 
2: for  $e \in M$  do
3:    $A_{\text{setup}}^e(M, w)$  outputs  $I', \mathbf{x}$ 
4:   For  $i \in S$ :  $\mathbf{x} = A_{\text{update}}^e(I', i, \mathbf{x})$ 
5:   If  $A_{\text{final}}^e(I', \mathbf{x}) = 1$ ,  $S' = S' \cup \{e\}$ 
6: end for
7: Return  $S'$ 

```

final values of \mathbf{x} associated with the partition that S belongs to, i.e., $\mathbf{x}_{S, \text{final}}$; let $m_i(\mathbf{v})$ be the agent i 's marginal contribution to subsets associated with $\mathbf{v} \in \mathcal{X}$. By Lemma 1,

$$\phi_i = \sum_{\mathbf{v} \in \mathcal{X}} \sum_{s=0}^{|N|-1} \hat{c}(|N| - 1, s, \mathbf{v}) \frac{s!(|N| - s - 1)!}{|N|!} m_i(\mathbf{v}),$$

and the Shapley value can be calculated in time $O((t_{\text{update}} + t_{\text{final}})m^n|N|)$ using the precomputed values of \hat{c} . The overall running time is $O(t_{\text{setup}} + t_{\text{update}}m^n|N|^2 + (t_{\text{update}} + t_{\text{final}})m^n|N|)$.

Now assume that the agents have to be processed in a specific order determined by representation I . For a given S and its final values $\mathbf{x}_{S, \text{final}}$, we cannot compute $\mathbf{x}_{S \cup \{i\}, \text{final}}$ as $A_{\text{update}}^e(I', i, \mathbf{x}_{S, \text{final}})$ and compute agent i 's marginal contribution to S , because it would violate the order if some agents in S have to be processed after i . Instead, we associate S with the final values $\mathbf{x}_{S, \text{final}}$ and $\mathbf{x}_{S \cup \{i\}, \text{final}}$ and partition $2^{N \setminus \{i\}}$ by the tuple $(|S|, \mathbf{x}_{S, \text{final}}, \mathbf{x}_{S \cup \{i\}, \text{final}})$ into at most $m^{2n}|N|$ partitions, omitting the argument I . Following the same argument as before, we get the running time $O(t_{\text{setup}} + t_{\text{update}}m^{2n}|N|^2 + t_{\text{final}}m^{2n}|N|)$. \square

The following definition and theorem generalize Theorem 4 and can also be considered a specialization of Theorem 5. See Appendix C for proof details.

Definition 2. Assume a cooperative game (N, v) has an alternative representation (M, w, A) as described in Lemma 3 such that $v(S) = \sum_{e \in A(S)} w(e)$, for all $S \subseteq N$. Algorithm A has a per-element decomposition $(A_{\text{setup}}^e, A_{\text{update}}^e, A_{\text{final}}^e)$ for all $e \in M$ if $A(S)$ can be computed as in Algorithm 3. We denote the upper bounds, over all $e \in M$, on running times of the sub-algorithms of the per-element decomposition t_{setup} , t_{update} and t_{final} , respectively.

Theorem 6. Assume a cooperative game (N, v) has an alternative representation (M, w, A) , as given in Lemma 3. If A has a per-element decomposition $(A_{\text{setup}}^e, A_{\text{update}}^e, A_{\text{final}}^e)$ for all $e \in M$ such that at most $n(M, w)$ variables \mathbf{x} are used with each taking at most $m(M, w)$ possible values as S ranges over all subsets of N and e over M , the Shapley value can be computed in $O((t_{\text{setup}} + t_{\text{update}}m^n|N|^2 + t_{\text{final}}m^n|N|)|M|)$ for each agent. In order-specific cases, for Step 4 of Algorithm 3, the running time is $O((t_{\text{setup}} + t_{\text{update}}m^{2n}|N|^2 + t_{\text{final}}m^{2n}|N|)|M|)$. Note that n and m are representation-dependent numbers and the argument (M, w) has been omitted.

The above definitions apply broadly and suggest the following framework for cooperative games that we term *algorithmic representation*; we represent each cooperative game (N, v) in terms of an alternative representation I and an accompanying algorithm A that computes v . As we can represent any cooperative game by a table with exponentially many entries for v values and a simple lookup algorithm, the algorithmic representation always exist. The main challenge is to determine an “efficient” algorithmic representation for cooperative games in general. The algorithmic representation framework subsumes the concise representation schemes in [6, 10, 1] as these assume specific structures on the alternative representation I . It also captures the notion of classes of cooperative games for we can represent a class of cooperative games by a set of alternative representations corresponding to those games in the class. In this framework, Theorems 5 and 6 show that if the algorithms for computing v satisfy the decomposability properties outlined in Definitions 1 and 2, then the Shapley value can be computed efficiently as long as these algorithms are efficient.

Using the generalizations, we can reproduce many previous results on efficient computation of the Shapley value up to a (pseudo) polynomial factor in the running time.⁸ As concrete examples, we prove several such results (and a new one on the Data Mining Problem in Section 1). We defer proofs to Appendix C:

Corollary 1. (*Weighted Majority Games*) Assume a cooperative game (N, v) has a representation given by $|N| + 1$ nonnegative integers $q, w_1, \dots, w_{|N|}$ such that $v(S)$ is 1 if $\sum_{i \in S} w_i \geq q$, or 0 otherwise. Then, the Shapley value can be computed in pseudo-polynomial time $O(q|N|^2)$ for each agent. (Identical to [14])

Corollary 2. (*MC-net Representation*) Assume a cooperative game (N, v) has a marginal-contribution (MC) net representation with boolean rules $R = \{r_1, \dots, r_m\}$ with each r_i having value v_i and of the form $(p_1 \wedge \dots \wedge p_a \wedge \neg n_1 \wedge \dots \wedge \neg n_b)$ such that $v(S) = \sum_{r_i: S \text{ satisfies } r_i} v_i$ for all $S \subseteq N$.⁹ Then, the Shapley value can be computed in $O(m|N|^2(\max_i |r_i|)^2)$ for each agent, where $|r|$ is the number of literals in rule r . (Compare to $O(m \max_i |r_i|)$, linear time in the representation size, in [10])

Corollary 3. (*Multi-Issue Representation*) Assume a cooperative game (N, v) has a multi-issue representation with subsets $C_1, \dots, C_t \subseteq N$ and characteristic functions $v_i : 2^{C_i} \rightarrow \mathbb{R}$ for all i such that $v(S) = \sum_{i=1}^t v_i(S \cap C_i)$ for all $S \subseteq N$. Then, the Shapley value can be computed in $O(t2^{\max_i |C_i|}|N|^2 \max_i |C_i|)$ for each agent. (Compare to $O(t2^{\max_i |C_i|})$ in [6])

Corollary 4. (*Data Mining Problem*) Assume a cooperative game (N, v) has a representation given by $|N| + 1$ nonnegative integers $k, w_1, \dots, w_{|N|}$ such that

⁸ The slightly slower running times can be attributed to our generalizations’ inability to derive closed form expressions on a game-by-game basis; for instance, evaluating the sum $\sum_{i=1}^n i$ in $O(n)$ instead of using the identity $\frac{n(n+1)}{2} = \sum_{i=1}^n i$ in $O(1)$. As generalizations apply in a black-box manner, we argue the loss in running time is reasonable for (pseudo) polynomial time computation.

⁹ If $r = (1 \wedge 2 \wedge \neg 3)$, then $S = \{1, 2\}$ satisfies r , but $S = \{1, 3\}$ does not.

$v(S) = \max_{S' \subseteq S: |S'| \leq k} w(S')$. Then, the Shapley value can be computed in polynomial time $O(|N|^3)$ for each agent. (This is our own problem.)

7 Further Discussion

We have introduced a class of cooperative games called budgeted games and investigated the computational complexity of the Shapley value in the knapsack version, knapsack budgeted games, in particular. We presented exact and approximation algorithms for knapsack budgeted games and a pseudo-polynomial time algorithm for closely related greedy knapsack budgeted games. These algorithms have only polynomial dependence on $|N|$, the number of agents, and are more efficient than the naïve exponential time algorithm when $|N|$ is large. Our results extend to knapsack budgeted games with multiple budget constraints. We believe knapsack budgeted games are useful in modeling value division problems in real-life scenarios and our algorithms applicable; for example, when finding a profit distribution solution for a joint venture of, say, 100-plus agents.

We also provided generalizations and proposed the algorithmic representation framework in which we represent each cooperative game in terms of an alternative representation and an accompanying algorithm that computes the underlying value function. We formalized efficient algorithmic representations and used the generalizations to show that computing the Shapley value in those cooperative games with efficient algorithmic representations can be done efficiently. To demonstrate the generalizations' applicability, we proved old and new results on the efficient computation of the Shapley value.

We note that further improvement to our algorithmic results might be possible. While the exact algorithm in Theorem 2 has polynomial time dependence on $|N|$, it is not a pseudo-polynomial time algorithm and the hardness result in Theorem 1 does not preclude the existence of a polynomial time algorithm for the Shapley value computation in the restricted case of $|N| \gg l_{\text{bin}}$.¹⁰ Similarly, we do not know if the results in Theorems 3 and 4 are the best possible. We pose these as open problems.

Finally, we believe our techniques can have applications beyond the games considered in this paper and to other economic concepts such as the Banzhaf index. It would be also interesting to investigate the computational complexity of the Shapley value in other kinds of budgeted games.

Acknowledgements. We would like to thank Vasilis Gkatzelis for his helpful comments.

References

- [1] K. V. Aadithya, T. P. Michalak, and N. R. Jennings. Representation of coalitional games with algebraic decision diagrams. AAMAS'11.

¹⁰ In this case, the $O(|N|l_{\text{bin}})$ dynamic programming time algorithm for the classical knapsack problem algorithm in Footnote 6 becomes an $O(|N|^2)$ algorithm, and the classical knapsack problem can be solved in polynomial time.

- [2] H. Aziz and T. B. Sorensen. Path coalitional games. *CoopMAS'11*.
- [3] Y. Bachrach, O. Lev, S. Lovett, J. S. Rosenschein, and M. Zadimoghaddam. Co-operative weakest link games. *AAMAS'14*. *To Appear*.
- [4] Y. Bachrach, E. Markakis, E. Resnick, A. D. Procaccia, J. S. Rosenschein, and A. Saberi. Approximating power indices: Theoretical and empirical analysis. *Autonomous Agents and Multi-Agent Systems*, Mar. 2010.
- [5] Y. Bachrach and E. Porat. Path disruption games. *AAMAS'10*.
- [6] V. Conitzer and T. Sandholm. Computing shapley values, manipulating value division schemes, and checking core membership in multi-issue domains. *AAAI'04*.
- [7] X. Deng and C. H. Papadimitriou. On the complexity of cooperative solution concepts. *Mathematics of Operations Research*, 19(2), 1994.
- [8] U. Faigle and W. Kern. On some approximately balanced combinatorial cooperative games. *Zeitschrift fr Operations Research*, 38(2), 1993.
- [9] S. S. Fatima, M. Wooldridge, and N. R. Jennings. A linear approximation method for the shapley value. *Artificial Intelligence*, 172(14), 2008.
- [10] S. Jeong and Y. Shoham. Marginal contribution nets: A compact representation scheme for coalitional games. *EC '05*.
- [11] S. Jeong and Y. Shoham. Multi-attribute coalitional games. *EC '06*.
- [12] J. Kuipers. Bin packing games. *Mathematical Methods of Operations Research*, 47(3), 1998.
- [13] R. T. Ma, D. Chiu, J. C. Lui, V. Misra, and D. Rubenstein. Internet economics: The use of shapley value for isp settlement. *CoNEXT'07*.
- [14] T. Matsui and Y. Matsui. A survey of algorithms for calculating power indices of weighted majority games. *J. Oper. Res. Soc. Japan*, 2000.
- [15] Y. Matsui and T. Matsui. Np-completeness for calculating power indices of weighted majority games. *Theoretical Computer Science*, 2001.
- [16] T. P. Michalak, K. V. Aadithya, P. L. Szczepanski, B. Ravindran, and N. R. Jennings. Efficient computation of the shapley value for game-theoretic network centrality. *J. Artif. Int. Res.*, Jan. 2013.
- [17] V. Misra, S. Ioannidis, A. Chaintreau, and L. Massoulié. Incentivizing peer-assisted services: A fluid shapley value approach. *SIGMETRICS '10*.
- [18] R. Narayanam and Y. Narahari. A shapley value-based approach to discover influential nodes in social networks. *IEEE Transactions on Automation Science and Engineering*, 8(1):130–147, Jan 2011.
- [19] X. Qiu. Bin packing games. Master's thesis, University of Twente, 2010.
- [20] L. S. Shapley. A value for n-person games. *Contributions to the theory of games*, 2:307–317, 1953.
- [21] V. V. Vazirani. *Approximation Algorithms*. Springer-Verlag New York, Inc., New York, NY, USA, 2001.

A Missing Proofs from Section 4

Proof. (Theorem 1) We reduce the decision version of the classical knapsack problem, which is NP-complete, to the problem of computing the Shapley value. If there is a polynomial time algorithm for the Shapley value computation, we use it to compute the Shapley value of the agents, $\phi_1, \dots, \phi_{|N|}$. By the efficiency property, $\phi_1 + \dots + \phi_{|N|} = v(N)$, which is exactly the solution of the optimization version of the classical knapsack problem. Then, we can solve the decision version. \square

Proof. (Lemma 1) We start from (1):

$$\begin{aligned}
\phi_i &= \sum_{S \subseteq N \setminus \{i\}} f(|S|)(v(S \cup \{i\}) - v(S)) \\
&= \sum_{S \subseteq N \setminus \{i\}} \sum_{p=1}^{p_i} \sum_{s=0}^{|N|-1} f(s) m_i(p) \text{Id}(|S| = s, P_i(S) = p) \\
&= \sum_{p=1}^{p_i} \sum_{s=0}^{|N|-1} f(s) m_i(p) \left(\sum_{S \subseteq N \setminus \{i\}} \text{Id}(|S| = s, P_i(S) = p) \right) \\
&= \sum_{p=1}^{p_i} \sum_{s=0}^{|N|-1} c(i, p, s) \frac{s!(N-s-1)!}{N!} m_i(p),
\end{aligned}$$

where $f(x) = \frac{x!(|N|-x-1)!}{|N|!}$ for a nonnegative integer x . Given the Shapley value equation, the running time is straightforward to obtain. \square

Proof. (Lemma 2) We bound the agent i 's marginal contribution to S with respect to v' : $v'(S \cup \{i\}) - v'(S) \leq v(S \cup \{i\}) - (v(S) - \alpha) = v(S \cup \{i\}) - v(S) + \alpha$, and $v'(S \cup \{i\}) - v'(S) \geq v(S \cup \{i\}) - v(S) - \alpha$. Then, marginal contributions computed with respect to v and v' are within α of each other. Using the Shapley value equation $\phi_i = \frac{1}{|N|!} \sum_{\pi \in \Pi} v(P_\pi^i \cup \{i\}) - v(P_\pi^i)$,

$$\begin{aligned}
|\phi_i - \phi'_i| &= \left| \frac{1}{|N|!} \sum_{\pi \in \Pi} (v(P_\pi^i \cup \{i\}) - v(P_\pi^i)) - \frac{1}{|N|!} \sum_{\pi \in \Pi} (v'(P_\pi^i \cup \{i\}) - v'(P_\pi^i)) \right| \\
&\leq \frac{1}{|N|!} \sum_{\pi \in \Pi} |(v(P_\pi^i \cup \{i\}) - v(P_\pi^i)) - (v'(P_\pi^i \cup \{i\}) - v'(P_\pi^i))| \\
&\leq \frac{1}{|N|!} \sum_{\pi \in \Pi} \alpha \\
&= \alpha.
\end{aligned}$$

\square

B Missing Proofs from Section 5

Proof. (Lemma 3) The proof is nearly identical to that of Lemma 1. We start from (1):

$$\begin{aligned}
\phi_i &= \sum_{S \subseteq N \setminus \{i\}} f(|S|)(v(S \cup \{i\}) - v(S)) \\
&= \sum_{S \subseteq N \setminus \{i\}} f(|S|) \left(\sum_{e \in A(S \cup \{i\})} w(e) - \sum_{e \in A(S)} w(e) \right) \\
&= \sum_{S \subseteq N \setminus \{i\}} \sum_{e \in M} \sum_{s=0}^{|N|-1} f(s) w(e) (\text{Id}(C_s, e \in A(S \cup \{i\})) - \text{Id}(C_s, e \in A(S))) \\
&= \sum_{e \in M} \sum_{s=0}^{|N|-1} f(s) w(e) \left(\sum_{S \subseteq N \setminus \{i\}} \text{Id}(C_s, e \in A(S \cup \{i\})) - \sum_{S \subseteq N \setminus \{i\}} \text{Id}(C_s, e \in A(S)) \right) \\
&= \sum_{e \in M} \sum_{s=0}^{|N|-1} (c_+(i, e, s) - c_-(i, e, s)) \frac{s!(|N| - s - 1)!}{|N|!} w(e),
\end{aligned}$$

where $f(x) = \frac{x!(|N| - x - 1)!}{|N|!}$ for a nonnegative integer x and C_s is the clause $|S| = s$.

Given the Shapley value equation, the running time is straightforward to obtain. \square

Proof. (Theorem 4) We compute the Shapley value ϕ_i for some fixed agent i . In what follows, we assume that the agents are sorted and reindexed so that $\frac{w_1}{l_1} \geq \dots \geq \frac{w_{|N|}}{l_{|N|}}$. If there are multiple agents with the same maximum weight for the argmax operator, we choose the one with the lowest index.

For ease of exposition, we use \mathcal{A}'' to denote Step 2 of the greedy heuristic so that $\mathcal{A}''(S, b)$ is exactly the set S' in $\mathcal{A}'(S, b)$. We also drop the bin size b when it is equal to l_{bin} . Note $v(S) = \sum_{k \in \mathcal{A}'(S)} w_k = w(\mathcal{A}'(S))$ for all $S \subseteq N$. In order to use Lemma 3, we choose the alternative representation (M, w, A) where $M = N$, the weight function w is such that $w(e) = w_e, \forall e$, and A is the greedy heuristic \mathcal{A}' . We consider three cases: $e = i$, $e > i$, and $e < i$.

Case 1) $e = i$: For $c_-(i, s, i)$, we count subsets $S \subseteq N \setminus \{i\}$ such that $i \in \mathcal{A}'(S)$. Clearly, $c_-(i, s, i) = 0$. For $c_+(i, s, i)$, we count subsets $S \subseteq N \setminus \{i\}$ such that $i \in \mathcal{A}'(S \cup \{i\})$. Note $i \in \mathcal{A}'(S \cup \{i\})$ if 1) agent i is included in S' in Step 2 of \mathcal{A}' and S' is finally returned in Step 3, or 2) agent i is not included in S' , but is selected in Step 1 and finally returned in Step 3. To each subset $S \subseteq N \setminus \{i\}$, we divide the set S into two parts, $S_{1,i-1}$ and $S_{i+1,|N|}$, and associate tuple $(s_1, \hat{l}, \hat{w}_1, a_1, s_2, \hat{w}_2, a_2)$ where $|S_{1,i-1}| = s_1$, $l(\mathcal{A}''(S_{1,i-1})) = \hat{l}$, $w(\mathcal{A}''(S_{1,i-1})) = \hat{w}_1$, $\text{argmax}_{k \in S_{1,i-1}} w_k = a_1$, $|S_{i+1,|N|}| = s_2$, $w(\mathcal{A}''(S_{i+1,|N|})) = \hat{w}_2$, and $\text{argmax}_{k \in S_{i+1,|N|}} w_k = a_2$. Note the tuple has enough information to reconstruct the steps of the greedy heuristic and check aforementioned conditions.

Since the total number of possible tuples is bounded, we count the number of sets $S \subseteq N \setminus \{i\}$ for each tuple and sum up to determine $c_+(i, s, i)$. We further decouple steps of \mathcal{A}' on $S_{1,i-1}$ and $S_{i+1,|N|}$ by using intermediate bin sizes, the parameter b in the coefficient \hat{c}_b^2 below, and consider each part independently using smaller tuples $(s_1, \hat{l}, \hat{w}_1, a_1)$ and (s_2, \hat{w}_2, a_2) , respectively.

We define $\hat{c}^1(j, s_1, \hat{l}, \hat{w}_1, a_1) = \#\{S \subseteq N_{1,j} : |S| = s_1, l(\mathcal{A}''(S)) = \hat{l}, w(\mathcal{A}''(S)) = \hat{w}_1, \arg\max_{k \in S} w_k = a_1\}$, for $1 \leq j \leq i-1$, and $\hat{c}_b^2(j, s_2, \hat{w}_2, a_2) = \#\{S \subseteq N_{i+1,j} : |S| = s_2, w(\mathcal{A}''(S, b)) = \hat{w}_2, \arg\max_{k \in S} w_k = a_2\}$, for $i+1 \leq j \leq |N|$ and $0 \leq b \leq l_{\text{bin}}$. Then, the number of subsets $S \subseteq N \setminus \{i\}$ with tuple $(s_1, \hat{l}, \hat{w}_1, a_1, s_2, \hat{w}_2, a_2)$ with i is contained in S' in Step 2 of $\mathcal{A}'(S)$ is the product of the number of subsets $S_1 \subseteq N_{1,i-1}$ with tuple $(s_1, \hat{l}, \hat{w}_1, a_1)$ and the number of subsets $S_2 \subseteq N_{i+1,|N|}$ with tuple (s_2, \hat{w}_2, a_2) , where the effective bin size for S_2 is $l_{\text{bin}} - \hat{l} - l_i$. These sets satisfy the first set of conditions for $i \in \mathcal{A}'(S \cup \{i\})$ if and only if $\text{Id}(\hat{l} \leq l_{\text{bin}} - l_i, \hat{w}_1 + \hat{w}_2 + w_i \geq \max\{w_{a_1}, w_{a_2}\}) = 1$. We reason similarly for the second set of conditions and get $c_+(i, s, i) = \sum \left(\hat{c}^1 \cdot \hat{c}_{l_{\text{bin}} - \hat{l} - l_i}^2 \cdot \text{Id}(C_1, C_2) + \hat{c}^1 \cdot \hat{c}_{l_{\text{bin}} - \hat{l}}^2 \cdot \text{Id}(C_3, C_4, C_5) \right)$, where the summation is over $s_1 + s_2 = s$, $1 \leq a_1 < a_2 \leq |N|$, $0 \leq \hat{l} \leq l_{\text{bin}}$, $0 \leq \hat{w}_1, \hat{w}_2 \leq w_{\text{max}}|N|$; \hat{c}^1 's have argument $(i-1, s_1, \hat{l}, \hat{w}_1, a_1)$ and \hat{c}^2 's have argument $(|N|, s_2, \hat{w}_2, a_2)$; the conditions are $C_1 = (\hat{l} \leq l_{\text{bin}} - l_i)$, $C_2 = (\hat{w}_1 + \hat{w}_2 + w_i \geq \max\{w_{a_1}, w_{a_2}\})$, $C_3 = (\hat{l} > l_{\text{bin}} - l_i)$, $C_4 = (i = \arg\max_{k \in \{a_1, a_2, i\}} w_k)$, and $C_5 = (w_i > \hat{w}_1 + \hat{w}_2)$. Using dynamic programming, \hat{c}^1 can be computed in $O(l_{\text{bin}} w_{\text{max}} |N|^3)$ and \hat{c}_b^2 for all $0 \leq b \leq l_{\text{bin}}$ can be computed in $O(l_{\text{bin}} w_{\text{max}} |N|^3)$. Hence, each coefficient $c_+(i, s, i)$ can be computed in $O(l_{\text{bin}} w_{\text{max}}^2 |N|^4)$ using precomputed values of \hat{c}^1 and \hat{c}_b^2 .

The analyses for the other cases are similar. We divide the set N at i and e into three parts and associate each subset satisfying $e \in \mathcal{A}'(S \cup \{i\})$ or $e \in \mathcal{A}'(S)$ a tuple that summarizes the steps of the greedy heuristic over these parts. Then, we proceed as before and compute each coefficient, $c_+(i, s, e)$ and $c_-(i, s, e)$, in $O(l_{\text{bin}}^5 w_{\text{max}}^5 |N|^6)$.

Putting all the cases together, we can compute the necessary coefficients $c(i, e, s)$ in time $O(l_{\text{bin}}^5 w_{\text{max}}^5 |N|^6)$ per coefficient. By Lemma 3, we can compute the Shapley value in pseudo-polynomial time of $O(l_{\text{bin}}^5 w_{\text{max}}^5 |N|^8)$. \square

C Missing Proofs from Section 6

Proof. (Theorem 6) Given the alternative representation (M, w) , we compute the Shapley value of agent i using Lemma 3. We show how to compute the quantity $\phi_{i,e} = \sum_{s=0}^{|N|-1} (c_+(i, s, e) - c_-(i, s, e)) \frac{s!(|N|-s-1)!}{|N|!}$ for some arbitrary $e \in M$. We associate $v(S)$ with the final states, $\mathbf{x}_{S, \text{final}}$, of $n(M, w)$ variables \mathbf{x} in A for all $S \subseteq N \setminus \{i\}$. We partition $2^{N \setminus \{i\}}$ by the pair $(|S|, \mathbf{x}_{S, \text{final}})$ into at most $m^n |N|$ partitions, omitting the argument (M, w) from n and m . Let \mathcal{X} be the set of all final values of the variables \mathbf{x} in A . We compute the cardinalities of the partitions using dynamic programming. Let $N' = N \setminus \{i\}$, ordered and

reabeled $1, \dots, |N| - 1$ and $i = |N|$. For $0 \leq j \leq |N| - 1$, $0 \leq s \leq j$, and $\mathbf{v} \in \mathcal{X}$, we define $\hat{c}(j, s, \mathbf{v}) = \#\{S \subseteq N'_{1,j} : |S| = s, \mathbf{x}_{S, \text{final}} = \mathbf{v}\}$. \hat{c} has the recurrence relation $\hat{c}(j, s, \mathbf{v}) = \hat{c}(j - 1, s, \mathbf{v}) + \sum_{\mathbf{u}: \mathcal{A}_{\text{update}}^e(I', j, \mathbf{u}) = \mathbf{v}} \hat{c}(j - 1, s - 1, \mathbf{u})$ with the base case $\hat{c}(0, 0, \mathbf{s}) = 1$, where \mathbf{s} is the initial states of variables \mathbf{x} . Using A_{setup}^e , we compute I' and the initial values \mathbf{s} in $O(t_{\text{setup}})$. Using the recurrence relation and $\mathcal{A}_{\text{update}}^e$, we compute $\hat{c}(j, s, \mathbf{v})$ for all j, s , and \mathbf{v} in time $O(t_{\text{update}} m^n |N|^2)$. Note that for a subset $S \subseteq N \setminus \{i\}$, we can compute if $e \in A(S)$ and if $e \in A(S \cup \{i\})$ in $O(t_{\text{final}})$ using the information associated with the partition that S belongs to. Then, we can compute the quantity $\phi_{i,e}$ in $O(t_{\text{final}} m^n |N|)$ using precomputed values of \hat{c} , and the Shapley value $\phi_i = \sum_{e \in M} \phi_{i,e}$ in $O((t_{\text{setup}} + t_{\text{update}} m^n |N|^2 + t_{\text{final}} m^n |N|)|M|)$ overall.

For the order-specific case, we follow the same line of reasoning as in Theorem 5. \square

Proof. (Corollary 1) We use the order-agnostic version of Theorem 5 with representation $I = (q, w_1, \dots, w_{|N|})$. We construct algorithm A with decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$ as follows. As variables \mathbf{x} , we use an indicator variable taking values $0, \dots, q$ to represent that total weight of subset of agents, capped at q . In A_{setup} , we initialize \mathbf{x} to 0. In A_{update} , for given $i \in S$, we update the indicator variable by adding w_i to its value while capping the total at q . In A_{final} , we return 1 if $\mathbf{x} = q$, or 0 otherwise. Therefore, $n(I) = 1$, $m(I) = q + 1$, $t_{\text{setup}} = O(1)$, $t_{\text{update}} = O(1)$, and $t_{\text{final}} = O(1)$. This leads to an $O(q|N|^2)$ algorithm for computing the Shapley value for each agent. \square

Proof. (Corollary 2) We use the order-agnostic version of Theorem 6 with representation (M, w, A) where $M = \{e_r : r \in R\}$, $w(e_r) = v_r$ for each rule r , and $A(S) = \{e_r : S \text{ satisfies } r\}$. For each $e_r \in M$, we construct a per-element decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$ for A , omitting the superscript e . Assume the rule r is $(p_1 \wedge \dots \wedge p_a \wedge \neg n_1 \wedge \dots \wedge \neg n_b)$ where each literal represents an agent. In A_{setup} , we create a $0/\pm 1$ array of length $|N|$ with $+1$'s in the positions corresponding to agents with positive literals in r , -1 's in the positions corresponding to agents with negative literals, and 0 's elsewhere. As variables \mathbf{x} , we use two counters: x_1 that is initialized to 0 and incremented by 1 for each occurrence of positive literal agent in S , and x_2 that is initialized to b and decremented by 1 for each occurrence of negative literal agent in S . In words, x_1 is equal to the number of positive literal agents in S and $b - x_2$ is equal to the number of negative literal agents in S . In A_{update} , for given $i \in S$, we check the agent's value in the array and update the variables accordingly. In A_{final} , we output 1 if $x_1 = a$ and $x_2 = b$, or 0 otherwise. Therefore, $n(M, w) = 2$, $m(M, w) = \max_i |r_i|$, $t_{\text{setup}} = O(|N|)$, $t_{\text{update}} = O(1)$, and $t_{\text{final}} = O(1)$. This leads to an $O(m|N|^2(\max_i |r_i|)^2)$ algorithm for computing the Shapley value for each agent. \square

Proof. (Corollary 3) We use the order-agnostic version of Theorem 6 with representation (M, w, A) where $M = \{e_{iC} : 1 \leq i \leq t, C \subseteq C_i\}$, $w(e_{iC}) = v_i(C)$ for all i and $C \subseteq C_i$, and $A(S) = \{e_{1, S \cap C_1}, \dots, e_{t, S \cap C_t}\}$ for all $S \subseteq N$. Note

$|M| = \sum_i 2^{|C_i|}$. We construct per-element decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$ for each $e \in M$ for A , omitting superscript e . Assume $e_{iC} \in M$. As variable \mathbf{x} , we use a counter for the number of elements of S , to be given as an input, in C . In A_{setup} , we initialize a 0/1 array of length $|N|$ with 1's in the positions corresponding to elements in C and 0's elsewhere. In A_{update} , for given $i \in S$, we increment \mathbf{x} if $i \in C$. In A_{final} , we check that $\mathbf{x} = |C|$. Therefore, $n(M, w) = 1$, $m(M, w) = \max_i |C_i|$, $t_{\text{setup}} = O(|N|)$, $t_{\text{update}} = O(1)$, and $t_{\text{final}} = O(1)$. This leads to an $O(t 2^{\max_i |C_i|} |N|^2 \max_i |C_i|)$ algorithm for computing the Shapley value for each agent. \square

Proof. (Corollary 4) We use the order-agnostic version of Theorem 6 with representation (M, w, A) where $M = N$, $w(i) = w_i$ for all i , and $A(S) = \{\text{at most } k \text{ top agents by weight in } S\}$. For simplicity, we assume the weights are all different; otherwise, we break ties consistently. For each $e \in M$, we construct a per-element decomposition $(A_{\text{setup}}, A_{\text{update}}, A_{\text{final}})$, omitting the superscript e . As variables \mathbf{x} , we use a counter that counts the number of agents with weight greater than agent e 's, capped at k . In A_{setup} , we initialize the counter \mathbf{x} to 0. In A_{update} , for given $i \in S$, we increment the counter if $w_i > w_e$, capping the counter at k . In A_{final} , we output 1 if $\mathbf{x} < k$, or 0 otherwise. Therefore, $n(M, w) = 1$, $m(M, w) = k + 1$, $t_{\text{setup}} = O(1)$, $t_{\text{update}} = O(1)$, and $t_{\text{final}} = O(1)$. This leads to an $O(k|N|^2)$ algorithm for computing the Shapley value for each agent.

If $k \geq |N|$, then $v(S) = \sum_{i \in S} w_i$ for all S and the Shapley value of agent i is simply w_i . So, we can compute the Shapley value in $O(1)$ for each agent after reading the representation. If $k < |N|$, then the above algorithm suffices and its running time reduces to $O(|N|^3)$. In either case, we get a polynomial time algorithm. \square